

PERANCANGAN PROGRAM APLIKASI MARKET BASKET ANALYSIS UNTUK MENDUKUNG PERSEDIAAN BARANG DENGAN METODE FUZZY C-COVERING

Zahedi; Charies Chandra

Mathematics & Statistics Department, School of Computer Science, Binus University
Jln. K.H. Syahdan No. 9, Palmerah, Jakarta Barat 11480
zahedizahedi@binus.ac.id, cherieliquids@yahoo.com

ABSTRACT

Any difficulty in analyzing sales transaction data is often faced by a company due to the huge number of sales transactions and the limited tools to process the data. It results in losses for the company since it is difficult to estimate the goods supply for subsequent sales. In this paper a data mining application is designed to analyze sales data. Market Basket Analysis is a data mining application that aims to determine most purchased or used products at once by the consumer. The process of Market Basket Analysis is to analyze consumer buying habits by finding associations among products purchased by different customers. The method used in Market Basket Analysis is a method of Fuzzy c-Covering, which is one method to classify the elements of a universal set into partitions of fuzzy sets. This study found that the value of support and confidence is part of the Market Basket Analysis, computed using the Fuzzy c-Covering. The higher the limit, the more selected the analytical results obtained are.

Keywords: market basket analysis, buying habits, fuzzy c-covering, fuzzy sets data mining

ABSTRAK

Kesulitan dalam menganalisis data transaksi penjualan sering dihadapi suatu perusahaan karena banyaknya jumlah data dan keterbatasan alat pengolah data. Hal tersebut mengakibatkan kesulitan memperkirakan persediaan barang untuk penjualan berikutnya sehingga membawa kerugian bagi perusahaan. Pada penelitian ini dirancang suatu aplikasi data mining untuk menganalisis data transaksi penjualan. Market basket analysis merupakan salah satu aplikasi data mining yang bertujuan untuk menentukan produk-produk yang paling sering dibeli atau digunakan oleh para konsumen. Proses market basket analysis ini adalah menganalisis buying habits konsumen dengan menemukan asosiasi antar produk-produk yang berbeda yang dibeli oleh konsumen menggunakan Fuzzy c-Covering, yaitu suatu metode untuk mengklasifikasikan elemen-elemen dari suatu himpunan universal menjadi partisi-partisi berupa fuzzy sets. Penelitian ini mendapatkan hasil bahwa nilai support dan confidence merupakan bagian dari market basket analysis yang dihitung dengan menggunakan metode Fuzzy c-Covering. Semakin tinggi nilai batasan, hasil analisis yang didapat akan lebih terseleksi.

Kata kunci: market basket analysis, buying habits, fuzzy c-covering, fuzzy sets data mining

PENDAHULUAN

Ada beberapa faktor yang dapat mendukung keunggulan sebuah perusahaan dibandingkan dengan perusahaan lain. Salah satu faktor yang penting adalah dalam hal persediaan barang yang memadai, tanpa kelebihan dan kekurangan. Jika perusahaan mempunyai jumlah persediaan barang yang lebih banyak dibandingkan dengan jumlah permintaan, dapat mengakibatkan kerugian biaya karena barang tersebut tidak habis terjual, terutama barang-barang yang mempunyai tanggal kadaluarsa. Sebaliknya jika perusahaan mempunyai jumlah persediaan barang yang lebih sedikit dibandingkan dengan jumlah permintaan, dapat mengakibatkan konsumen tersebut pergi karena barang yang ingin dibeli oleh konsumen tersebut tidak mencukupi (*opportunity loss*).

Oleh karena itu, untuk menyimpan persediaan barang yang sesuai dengan permintaan konsumen, perusahaan perlu melakukan suatu penelitian dan perencanaan yang dapat memprediksi permintaan konsumen. Sehingga perusahaan dapat memperoleh keuntungan yang optimal tanpa harus mengeluarkan biaya persediaan yang banyak. Komputer, sebagai alat bantu hitung untuk menganalisis persoalan dalam dunia bisnis dan pemasaran bukan merupakan hal yang asing lagi sekarang ini. Dengan bantuan komputer, analisis kuantitatif yang dihasilkan akan lebih cepat dan tepat dibandingkan dengan menggunakan kemampuan manusia. Situs amazon.com menjadi contoh kasus yang baik untuk melihat implementasi langsung dari analisis yang memanfaatkan komputer. Ketika kita berkunjung ke situs ini, kita sebagai pengunjung sekaligus pembeli akan melihat informasi mengenai kecenderungan beli konsumen yang terjadi di sana (Hoobs et al., 2003).

Market Basket Analysis atau *MBA*, merupakan salah satu tipe analisis data yang paling sering digunakan dalam dunia pemasaran (Megaputer, 2007). Tujuan dari *Market Basket Analysis* adalah untuk menentukan produk-produk apa saja yang paling sering dibeli atau digunakan sekaligus oleh para konsumen. Proses *Market Basket Analysis* ini adalah dengan menganalisis *buying habits* konsumen dengan menemukan asosiasi antar produk-produk yang berbeda yang diletakkan konsumen dalam *shopping basket*.

Istilah *Market Basket Analysis* sendiri datang dari kejadian yang sudah sangat umum terjadi di dalam pasar swalayan, yaitu ketika para konsumen memasukkan semua barang yang mereka beli ke dalam keranjang (*basket*) yang umumnya telah disediakan oleh pihak swalayan itu sendiri. Informasi mengenai produk-produk yang biasanya dibeli secara bersama-sama oleh para konsumen dapat memberikan “wawasan” tersendiri bagi para pengelola toko atau swalayan untuk menaikkan laba bisnisnya (Albion Research, 2007). Informasi-informasi atau pengetahuan seperti di atas tentunya tidak hanya bermanfaat di dalam lingkungan pemasaran untuk pasar swalayan saja. Beberapa bisnis yang bergerak di luar wilayah ini pun bisa menikmati manfaat dari adanya *Market Basket Analysis* ini. Sebut saja misalnya toko-toko virtual yang menjual produk-produknya secara *online*, bank-bank yang memberikan fasilitas layanan kartu kredit untuk para nasabahnya, perusahaan penyedia jasa asuransi, restoran *fast-food*, toko baju, toko buku, dan lain-lain.

Untuk beberapa kasus, pola dari barang-barang yang dibeli secara bersamaan oleh konsumen mudah untuk ditebak, misalnya susu dibeli bersamaan dengan roti. Namun, mungkin saja terdapat suatu pola pembelian barang-barang yang tidak pernah terpikirkan sebelumnya. Misalnya, pembelian minyak goreng dengan deterjen. Mungkin saja pola seperti ini tidak pernah terpikirkan sebelumnya karena minyak goreng dan deterjen tidak mempunyai hubungan sama sekali, baik sebagai barang pelengkap maupun barang pengganti. Hal ini mungkin tidak pernah terpikirkan sebelumnya sehingga tidak dapat diantisipasi jika terjadi sesuatu, seperti kekurangan stok deterjen misalnya. Inilah salah satu manfaat yang dapat diperoleh dari melakukan *Market Basket Analysis*. Dengan melakukan proses ini secara otomatis, seorang manajer tidak perlu mengalami kesulitan untuk menemukan pola barang-barang apa saja yang mungkin dibeli secara bersamaan.

Market Basket Analysis adalah teknik matematis yang biasa digunakan oleh marketing profesional untuk menyatakan kesamaan antara produk individu atau kelompok produk. *Market Basket Analysis* berkenaan dengan sekumpulan permasalahan bisnis yang berkaitan untuk mengetahui *point of sale* dari data transaksi.

Market Basket Analysis merupakan salah satu contoh penerapan *Association Rule*. Untuk menyampaikan ide mendasar dari *Market Basket Analysis*, dapat dilihat dari keranjang belanjaan yang berisi bermacam-macam barang yang dibeli oleh seseorang di sebuah supermarket. Keranjang ini berisi bermacam-macam barang seperti roti, susu, sereal, telur, mentega, gula, dan sebagainya. Sebuah keranjang memberitahukan kepada kita tentang apa saja yang dibeli oleh seorang konsumen dalam satu waktu. Sebuah daftar belanjaan yang lengkap yang diperoleh dari semua konsumen memberikan kita informasi yang sangat banyak, dan ini dapat menjelaskan barang-barang apa saja yang paling penting dari bisnis penjualan yaitu “apa barang yang dibeli oleh konsumen dan kapan”.

Setiap konsumen membeli seperangkat barang-barang yang berbeda, dalam jumlah yang berbeda, dan dalam waktu yang berbeda. *Market Basket Analysis* menggunakan informasi apa yang dibeli oleh konsumen untuk menyediakan tanda/informasi yaitu siapa mereka dan mengapa mereka melakukan pembelian tersebut? *Market Basket Analysis* menyediakan pengertian tentang barang dagangan dengan memberitahukan kepada kita produk-produk mana yang memungkinkan untuk dibeli secara bersamaan dan produk mana yang lebih disetujui untuk dipromosikan.

Market Basket Data adalah data transaksi yang menjelaskan tiga perbedaan entitas yang mendasar yaitu: *customers*, *orders* (pembelian), dan *items* (barang-barang).

Pengenalan konsumen setiap saat memungkinkan untuk dikenali secara cepat, seperti frekuensi pembelian yang dilakukan oleh konsumen. Tiga level dari *market basket data* yang penting secara cepat dapat memahami permintaan. Ada beberapa dasar pengukuran yaitu: (1) berapa rata-rata pembelian barang yang dilakukan konsumen?; (2) berapa rata-rata barang yang khusus setiap pembelian?; (3) berapa rata-rata barang setiap pembelian?; (4) untuk barang tertentu, proporsi konsumen apa yang telah membeli barang tersebut?; (5) untuk barang tertentu, berapa rata-rata dari pembelian setiap konsumen yang termasuk dalam barang tersebut?; (6) untuk barang tertentu, berapa rata-rata jumlah pembelian dalam suatu pembelian ketika barang tersebut dibeli?

METODE

Association Rule

Association Rule adalah bentuk jika “kejadian sebelumnya” kemudian “konsekuensinya”, (*IF antecedent, THEN consequent*). Bersamaan dengan perhitungan aturan *support* dan *confidence*. *Association Rule* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*. Contoh dari *Association Rule* dari analisis pembelian di suatu pasar swalayan adalah bisa diketahui berapa besar kemungkinan seorang konsumen membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

Salah satu contoh penerapan *Association Rule* adalah *Market Basket Analysis*. *Association Rule* menjadi terkenal karena aplikasinya untuk menganalisis isi keranjang belanja di pasar swalayan, sehingga *Association Rule* juga sering disebut dengan istilah *Market Basket Analysis*. *Association Rule* juga dikenal sebagai salah satu teknik *data mining* yang menjadi dasar dari berbagai teknik *data mining* lainnya.

Association Rule Mining

Association Rule Mining adalah suatu prosedur untuk mencari hubungan antar *item* dalam suatu *data set* yang ditentukan. (Han, Kamber, 2001). Dalam dunia bisnis lazim dikenal istilah *affinity analysis*. Tugas dari *association rule* adalah mencari aturan yang tidak mencakup hubungan antara dua atau lebih atribut. *Association rule* meliputi dua tahap (Ulmer, David, 2002): (1) mencari kombinasi yang paling sering terjadi dari suatu *itemset*; (2) mendefinisikan *condition* dan *result* (untuk *conditional association rule*). Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran ketertarikan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu. Umumnya ada dua ukuran, yaitu: (1) *support* – suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item/itemset* dari keseluruhan transaksi. Ukuran ini menentukan apakah suatu *item/itemset* layak untuk dicari *confidence*-nya (misal, dari keseluruhan transaksi yang ada, seberapa besar tingkat dominasi yang menunjukkan bahwa *item* A dan B dibeli bersamaan); (2) *confidence* – suatu ukuran yang menunjukkan hubungan antar dua *item* secara *conditional* (misal, seberapa sering *item* B dibeli jika orang membeli *item* A). Kedua ukuran ini nantinya berguna dalam menentukan *interesting association rules*, yaitu untuk dibandingkan dengan batasan (*threshold*) yang ditentukan oleh *user*. Batasan tersebut umumnya terdiri dari *min_support* dan *min_confidence*.

Metodologi dasar analisis asosiasi terbagi menjadi dua tahap: (1) analisis pola frekuensi tinggi – tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam *database*, yang dapat dirumuskan sebagai berikut:

$$\text{Support} = P(A \cap B) = \frac{\text{jumlah transaksi yang memuat A dan B}}{\text{total jumlah transaksi}}$$

(2) pembentukan aturan asosiatif – setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$ dengan menggunakan rumus berikut:

$$\text{Confidence} = P(B/A) = \frac{\text{Support}(A \cap B)}{P(A)}$$

Berikut ini adalah contoh dari *association rules* yang menggunakan database penjualan sederhana (Tabel 1). Diketahui bahwa jumlah transaksi yang memuat {roti, mentega} ada 4 (*support* 80%), sedangkan jumlah transaksi yang memuat {roti, mentega, susu} ada 2 (*support* 40%), transaksi yang memuat {buncis} hanya 1 (*support* 20%), dan lain-lain. Bila ditetapkan syarat minimum dari nilai *support* untuk *association rules* dalam contoh ini adalah 30%, diperoleh pola kombinasi yang memenuhi syarat minimum nilai *support* adalah sebagai berikut (Tabel 2).

Tabel 1
Contoh Database Penjualan

Kode Transaksi	Produk yang Terjual
001	pena, roti, mentega
002	roti, mentega, telur, susu
003	buncis, telur, susu
004	roti, mentega
005	roti, mentega, kecap, telur, susu

Tabel 2
Pola Kombinasi yang Memenuhi Syarat Minimum Nilai Support

Kombinasi Produk	Support
{roti}	80%
{mentega}	80%
{telur}	60%

{susu}	60%
{roti, mentega}	80%
{roti, telur}	40%
{roti, susu}	40%
{mentega, telur}	40%
{mentega, susu}	40%
{telur, susu}	60%
{roti, mentega, telur}	40%
{roti, mentega, susu}	40%
{roti, telur, susu}	40%
{mentega, telur, susu}	40%
{roti, mentega, telur, susu}	40%

Setelah semua pola kombinasi ditemukan, barulah dicari *association rules* yang memenuhi syarat minimum untuk *confidence*. Bila syarat minimum untuk *confidence* dari contoh di atas adalah 50% maka salah satu contoh *association rules* yang dapat ditemukan adalah sebagai berikut (Tabel 3).

Tabel 3
Association rules yang Memenuhi Syarat Minimum Nilai Confidence.

Association Rules	Support	Confidence
{roti} → {mentega}	80%	100%
{roti} → {telur}	40%	50%
{roti} → {susu}	40%	50%
{roti} → {mentega, telur}	40%	50%
{roti} → {mentega, susu}	40%	50%
{roti} → {telur, susu}	40%	50%
{roti} → {mentega, telur, susu}	40%	50%
{mentega} → {roti}	80%	100%
{mentega} → {telur}	40%	50%
{mentega} → {susu}	40%	50%
{mentega} → {roti, telur}	40%	50%
{mentega} → {roti, susu}	40%	50%
{mentega} → {telur, susu}	40%	50%
{mentega} → {roti, telur, susu}	40%	50%
{telur} → {roti}	40%	66,6%
{telur} → {mentega}	40%	66,6%
{telur} → {susu}	60%	100%
{telur} → {roti, mentega}	40%	66,6%
{telur} → {roti, susu}	40%	66,6%
{telur} → {mentega, susu}	40%	66,6%
{telur} → {roti, mentega, susu}	40%	66,6%
{susu} → {roti}	40%	66,6%
{susu} → {mentega}	40%	66,6%
{susu} → {telur}	60%	100%
{susu} → {roti, mentega}	40%	66,6%
{susu} → {roti, telur}	40%	66,6%
{susu} → {mentega, telur}	40%	66,6%
{susu} → {roti, mentega, telur}	40%	66,6%
{roti, mentega} → {telur}	40%	66,6%
{roti, mentega} → {susu}	40%	50%
{roti, mentega} → {telur, susu}	40%	50%
{roti, telur} → {mentega}	40%	100%
{roti, telur} → {susu}	40%	100%
{roti, telur} → {mentega, susu}	40%	100%

{roti, susu} → {mentega}	40%	100%
{roti, susu} → {telur}	40%	100%
{roti, susu} → {mentega, telur}	40%	100%
{mentega, telur} → {roti}	40%	100%
{mentega, telur} → {susu}	40%	100%
{mentega, telur} → {roti, susu}	40%	100%
{mentega, susu} → {roti}	40%	100%
{mentega, susu} → {telur}	40%	100%
{mentega, susu} → {roti, telur}	40%	100%
{telur, susu} → {roti}	40%	66,6%
{telur, susu} → {mentega}	40%	66,6%
{telur, susu} → {roti, mentega}	40%	66,6%
{roti, mentega, telur} → {susu}	40%	100%
{roti, mentega, susu} → {telur}	40%	100%
{roti, telur, susu} → {mentega}	40%	100%
{mentega, telur, susu} → {roti}	40%	100%

Lift/Improvement Ratio

Lift ratio adalah parameter penting selain *support* dan *confidence* dalam *association rule*. *Lift ratio* mengukur seberapa penting rule yang telah terbentuk berdasarkan nilai *support* dan *confidence*. *Lift ratio* merupakan nilai yang menunjukkan kevalidan proses transaksi dan memberikan informasi apakah benar produk A dibeli bersamaan dengan produk B. *Lift/Improvement Ratio* dapat dihitung dengan rumus:

$$\frac{\text{Support (A} \cap \text{B)}}{\text{Support (A)} * \text{Support (B)}}$$

Sebuah transaksi dikatakan valid jika mempunyai nilai *Lift/Improvement* lebih dari 1, yang berarti bahwa dalam transaksi tersebut, produk A dan B benar-benar dibeli secara bersamaan.

Fuzzy Logic

Fuzzy Logic dikembangkan untuk menyediakan fungsi, dan aturan-aturan matematis yang memperbolehkan input berupa bahasa yang alami. Menurut Ter Meulen (2001), bahasa alami (*natural language*) adalah bahasa yang digunakan secara umum oleh manusia dalam berkomunikasi secara lisan, ataupun tertulis. Sering juga dikenal untuk membangun bahasa pemrograman atau juga “bahasa” yang digunakan dalam logika yang formal, terutama dalam logika matematika.

Teori Himpunan Fuzzy akan memberikan jawaban terhadap suatu masalah yang mengandung ketidakpastian. Aplikasi logika fuzzy untuk mendukung keputusan semakin diperlukan ketika semakin banyak kondisi yang menuntut adanya keputusan yang tidak hanya bisa dijawab dengan ‘ya’ atau ‘tidak’. *Fuzzy Logic* memberikan rata-rata dari perhitungan angka, yang terletak antara nilai benar mutlak dan nilai salah mutlak, yang berupa range antara 0,0 dan 1,0. Dengan *Fuzzy Logic*, pengguna dimungkinkan untuk menghitung derajat keanggotaan dari sebuah data. *Fuzzy Logic* berurusan dengan kondisi yang tidak pasti, di mana benar dan salah tidak dapat ditentukan secara mutlak. Kebanyakan dari metode ini berbicara tentang keambiguan, yang mana bisa kita temukan dalam kehidupan kita sehari-hari. Berbeda dengan logika konvensional yang secara alami dapat dihubungkan dengan kondisi Boolean (benar/salah; I/O), *Fuzzy Logic* mencoba untuk menentukan daerah di mana kebenaran yang mutlak tidak dapat dicapai, begitu pula dengan kesalahan yang mutlak, yaitu daerah diantara kebenaran dan kesalahan yang mutlak tersebut. Pada beberapa kasus khusus, seperti nilai keanggotaan yang kemudian akan menjadi 0 atau 1, teori dasar tersebut akan identik dengan teori himpunan biasa, dan himpunan Fuzzy akan menjadi himpunan *crisp* biasa.

Logika Fuzzy adalah suatu cara yang tepat untuk memetakan suatu ruang input ke dalam suatu ruang output. Sebagai contoh: (1) manajer pergudangan mengatakan pada manajer produksi “Seberapa banyak persediaan barang pada akhir minggu ini?”, yang kemudian akan ditindak lanjuti dengan menetapkan jumlah barang yang harus diproduksi atau disiapkan esok hari; (2) pelayan sebuah rumah makan memberikan pelayanan terhadap tamu, kemudian tamu akan memberikan tip yang sesuai atas baik tidaknya pelayanan yang diberikan; (3) anda mengatakan pada saya seberapa sejuk ruangan yang anda inginkan, saya akan mengatur putaran kipas dari penyejuk ruangan yang bekerja pada ruangan ini.

Logika Fuzzy dikatakan sebagai logika baru yang lama, sebab ilmu tentang logika Fuzzy modern dan metodis baru yang ditemukan beberapa tahun yang lalu, padahal sebenarnya konsep tentang logika fuzzy itu sendiri sudah ada sejak lama. Konsep dari *Fuzzy Logic* diperkenalkan oleh Professor Lotfi A. Zadeh, di Barkley pada Universitas California (University of California) pada 1960an.

Ada beberapa alasan mengapa orang menggunakan logika Fuzzy, antara lain: (1) konsep logika Fuzzy mudah dimengerti. Konsep matematis yang mendasari penalaran Fuzzy sangat sederhana dan mudah dimengerti; (2) logika Fuzzy sangat fleksibel; (3) logika Fuzzy memiliki toleransi terhadap data-data yang tidak tepat; (4) logika Fuzzy mampu memodelkan fungsi-fungsi non-linear yang sangat kompleks; (5) logika Fuzzy dapat membangun dan mengaplikasikan pengalaman-pengalaman para pakar secara langsung tanpa harus melalui proses pelatihan; (6) logika Fuzzy dapat bekerja sama dengan teknik-teknik kendali secara konvensional; (7) logika Fuzzy didasarkan pada bahasa alami.

Fuzzy c-Covering

Fuzzy c-Covering merupakan salah satu metode yang dipakai untuk mengklasifikasikan elemen-elemen dari suatu himpunan universal menjadi partisi-partisi berupa *fuzzy sets*. *Fuzzy c-Covering* sendiri merupakan generalisasi dari metode *fuzzy c-partition* yang telah dikenal sebelumnya. *Fuzzy c-partition* dapat didefinisikan sebagai berikut (Klir, Yuan, 2001):

Misalkan $I = \{i_1, i_2, \dots, i_n\}$ adalah *domain* dari data. *Fuzzy c-partition* dari I adalah *fuzzy subset* atau *fuzzy classes* dari T , ditunjukkan oleh $T = \{t_1, t_2, \dots, t_c\}$, yang memenuhi:

$$\sum_{m=1}^c \mu_{t_m}(i_k) = 1, \text{ untuk semua } k \in N_n \quad \dots (1)$$

$$0 < \sum_{k=1}^n \mu_{t_m}(i_k) < n, \text{ untuk semua } m \in N_c \quad \dots (2)$$

di mana: c adalah *positive integer* ($0 < c \leq n$) dan $\mu_{t_m}(i_k) \in [0,1]$.

Berikut ini contoh untuk memperjelas rumus 1 dan 2.

Contoh 1:

Misalkan ada suatu *fuzzy classes* sebagai berikut:

$$\begin{aligned} \mu_{t_1} &= \{1/i_1, 0.8/i_2, 0.2/i_3\}, \\ \mu_{t_2} &= \{0.2/i_2, 0.8/i_3, 0.4/i_4\}, \\ \mu_{t_3} &= \{0.6/i_4, 1/i_5, 1/i_6\}. \end{aligned}$$

Dari contoh 1 dapat diketahui jumlah elemen yang terlibat ($n = 6$) dan jumlah partisinya ($c = 3$). *Fuzzy class* tersebut dapat dikatakan sebagai *fuzzy c-partition* dari suatu *domain* $I = \{i_1, i_2, i_3, i_4, i_5, i_6\}$ karena memenuhi rumus (1) dan (2). Di dalam *Fuzzy c-Covering*, rumus (1) dari teori *fuzzy c-partition* (Intan, Mukaidono, 2003) digeneralisasi menjadi:

$$\sum_{m=1}^c \mu_{t_m}(i_k) \geq 1, \text{ untuk semua } k \in N_n \quad \dots (3)$$

Untuk memperjelasnya, diberikan contoh sebagai berikut:

Contoh 2

Misalkan ada suatu *fuzzy classes* sebagai berikut:

$$\mu_{t1} = \{1/i_1, 0.8/i_2, 0.4/i_3\},$$

$$\mu_{t2} = \{1/i_2, 0.8/i_3, 0.6/i_4, 0.2/i_6\},$$

$$\mu_{t3} = \{1/i_4, 1/i_5, 1/i_6\}.$$

Dari contoh 2 dapat diketahui $n = 6$ dan $c = 3$. Fuzzy class tersebut dapat dikatakan sebagai *Fuzzy c-Covering* dari suatu domain $I = \{i_1, i_2, i_3, i_4, i_5, i_6\}$ karena memenuhi rumus (2) dan (3).

Algoritma untuk Market Basket Analysis Berdasarkan Fuzzy c-Covering

Dalam mencari hubungan antar produk, metode *Fuzzy c-Covering* ini berdasarkan pada persepsi bahwa semakin banyak produk yang dibeli dalam satu transaksi, maka hubungan antar produk yang terdapat dalam transaksi itu semakin lemah.

Berikut adalah langkah-langkah dalam algoritma yang dibuat berdasarkan metode tersebut:

Langkah 1: menentukan *max_item_threshold* yang dibutuhkan. *Max_item_threshold* adalah suatu pembatas yang dipakai untuk menyaring transaksi berdasarkan jumlah produk dalam transaksi tersebut. Hal ini didasarkan atas pemahaman bahwa semakin banyak produk yang dibeli dalam suatu transaksi, hubungan antar produk dalam transaksi tersebut semakin lemah.

Langkah 2: mencari *record-record* dalam tabel transaksi yang memenuhi *max_item_threshold* dan menyimpannya ke dalam *QT*, di mana:

$$QT = \{t \mid |t| \leq ith, ith \in \text{positive integer}\}$$

di mana:

QT (Qualified Transaction) : himpunan transaksi yang memenuhi *max_item_threshold*;

t : transaksi;

|t| : jumlah produk dalam suatu transaksi;

ith : *max_item_threshold*.

Langkah 3: set $k = 1$ (k adalah variabel untuk menentukan jumlah kombinasi).

Langkah 4: menentukan *min_support* ke- k sebagai *threshold* bagi kombinasi *k-item* terhadap tingkat dominasinya dari keseluruhan transaksi.

Langkah 5: mencari *support* dari setiap kombinasi *k-item* yang memungkinkan yang ada di dalam transaksi tersebut dengan rumus:

$$\text{support}(u) = \frac{\sum_{t=1}^n \frac{1}{C_{|T_t|}^k s(u, T_t)}}{n} = \frac{\sum_{t=1}^n \frac{1}{\frac{|T_t|!}{k! (|T_t| - k)!} s(u, T_t)}}{n} \quad \dots (6)$$

di mana:

u : kombinasi *k-item* yang dicari *support*-nya. Jika I adalah *universal set of items*, $u \subseteq I$;

$|u| = k$: jumlah produk dalam u ;

T_t : transaksi ke- t ($T_t \subseteq I$);

$|T_t|$: jumlah produk dalam T_t .

$C_{|T_t|}^k$: kombinasi *k-item* terhadap $|T_t|$;

n : jumlah *record/tuple* dalam *QT*.

$s(u, T_i) \in \{0,1\}$ adalah suatu *function*, di mana: jika $u \in T_i$, $s(u, T_i) = 1$, selain itu $s(u, T_i) = 0$.

Langkah 6: melakukan penyaringan terhadap kombinasi produk yang ada di dalam transaksi tersebut yang tidak memenuhi: $support(u) \geq min_support$ ke- k .

Langkah 7: set $k = k + 1$, di mana jika $k > ith$, ke langkah 9.

Langkah 8: mencari kombinasi k -item yang memungkinkan dari tiap kombinasi $(k-1)$ -item yang memenuhi *minimum support* yang telah ditentukan, dengan cara: untuk mendapatkan kombinasi k -item, u , harus ada semua kombinasi $(k-1)$ -item, u' , di mana $u' \subset u$, misalnya untuk mendapatkan $u = \{I_1, I_2, I_3, I_4\}$, harus ada $u' = \{I_1, I_2, I_3\}$, $\{I_1, I_2, I_4\}$, $\{I_1, I_3, I_4\}$ dan $\{I_2, I_3, I_4\}$. Jika tidak ada lagi kombinasi k -item yang memungkinkan yang memenuhi *min_support* yang telah ditentukan, dapat ke langkah 9, selain itu ulangi langkah 4-7.

Langkah 9: mendefinisikan tiap produk yang telah didapat dari langkah-langkah di atas sebagai *fuzzy set* (disebut *item fuzzy set*) terhadap transaksi QT .

Langkah 10: mencari *candidate rules* dengan cara menghitung *confidence* dari setiap kombinasi k -item yang memenuhi *min_support* ke- k ($k \geq 2$) dari *item fuzzy set* yang telah didapat pada langkah 9 dengan rumus:

$$R(X, Y) = \text{confidence} \quad (Y \rightarrow X) = \frac{\sum_{t \in T} \inf_{i \in X \cup Y} (\mu_i(t))}{\sum_{t \in T} \inf_{i \in Y} (\mu_i(t))} \quad \dots (7)$$

di mana:

$X, Y \subseteq I$;

T : himpunan dari kode-kode transaksi yang ada dalam QT ;

$\mu_i(t) \in [0,1]$: fungsi anggota terhadap T .

HASIL DAN PEMBAHASAN

Berikut ini diberikan contoh aplikasi untuk memperjelas cara kerja *Market Basket Analysis* berdasarkan metode *Fuzzy c-Covering*. Misalkan ada transaksi seperti pada Tabel 4 berikut.

Tabel 4

Contoh Database Penjualan

Kode Transaksi	Produk yang Terjual
100	I1, I2, I5
200	I2, I4
300	I2, I3
400	I1, I2, I4
500	I1, I3
600	I2, I3
700	I1, I3
800	I1, I2, I3, I5
900	I1, I2, I3
1000	I1, I2, I4, I6, I8

Misalkan $max_item_threshold = 4$ dan $set\ k = 1$. Dari Tabel 4, yang memenuhi $max_item_threshold$ hanya transaksi berkode 100 s/d 900. Maka dari Tabel 4, $QT = \{(I1, I2, I5), (I2, I4), (I2, I3), (I1, I2, I4), (I1, I3), (I2, I3), (I1, I3), (I1, I2, I3, I5), (I1, I2, I3)\}$ dan $T = \{100, 200, 300, 400, 500, 600, 700, 800, 900\}$.

Kemudian tentukan $min_support$ ke-1. Misalkan $min_support_1 = 0,1 = 10\%$. Berdasarkan QT di atas, dapat ditentukan *item-item* yang dicari *support*-nya, yaitu I1, I2, I3, I4, dan I5. Berikut ini adalah cara penghitungan *support* untuk tiap *item*:

$$\begin{aligned}
 I1 &= \frac{\frac{1}{3} + 0 + 0 + \frac{1}{3} + \frac{1}{2} + 0 + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}}{9} = \frac{1}{4} = 0,25 \\
 I2 &= \frac{\frac{1}{3} + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + 0 + \frac{1}{2} + 0 + \frac{1}{4} + \frac{1}{3}}{9} = \frac{11}{36} = 0,306 \\
 I3 &= \frac{0 + 0 + \frac{1}{2} + 0 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}}{9} = \frac{31}{108} = 0,287 \\
 I4 &= \frac{0 + \frac{1}{2} + 0 + \frac{1}{3} + 0 + 0 + 0 + 0 + 0}{9} = \frac{5}{54} = 0,093 \quad (\text{tidak memenuhi}) \\
 I5 &= \frac{\frac{1}{3} + 0 + 0 + 0 + 0 + 0 + 0 + \frac{1}{4} + 0}{9} = \frac{7}{108} = 0,065 \quad (\text{tidak memenuhi})
 \end{aligned}$$

Dari hasil perhitungan di atas, yang memenuhi $min_support_1$ adalah I1, I2, dan I3. Setelah itu k di-set menjadi $k = 2$. Misalkan $min_support_2 = 9\%$. Kemudian dicari kombinasi 2-*item* dan *support* dari *item* yang tersisa:

$$\begin{aligned}
 \{I1, I2\} &= \frac{\frac{1}{3} + 0 + 0 + \frac{1}{3} + 0 + 0 + 0 + \frac{1}{6} + \frac{1}{3}}{9} = \frac{7}{54} = 0,13 \\
 \{I1, I3\} &= \frac{0 + 0 + 0 + 0 + 1 + 0 + 1 + \frac{1}{6} + \frac{1}{3}}{9} = \frac{5}{18} = 0,28 \\
 \{I2, I3\} &= \frac{0 + 0 + 1 + 0 + 0 + 1 + 0 + \frac{1}{6} + \frac{1}{3}}{9} = \frac{5}{18} = 0,28
 \end{aligned}$$

Seperti yang dapat dilihat, semua kandidat *itemset* di atas memenuhi $support(u) \geq min_support_2$, semua *itemset* tersebut dikombinasikan lagi menjadi kombinasi 3-*itemset*. Set $k=3$ dan misalkan $min_support_3 = 11\%$. *Support*-nya kemudian dicari sebagai berikut:

$$\{I1, I2, I3\} = \frac{0 + 0 + 0 + 0 + 0 + 0 + 0 + \frac{1}{4} + 1}{9} = \frac{5}{36} = 0,139$$

Karena sudah tidak ada kombinasi yang memungkinkan lagi untuk memenuhi $min_support$ yang ditentukan, maka penghitungan selesai. Setelah itu tiap *item* yang telah didapatkan dari langkah-langkah di atas, yaitu I1, I2, dan I3, didefinisikan sebagai *fuzzy set* terhadap T . Untuk lebih jelasnya, akan diberikan cara mendefinisikan *item* I1 sebagai *fuzzy set* terhadap $T = 100$ dan $T = 400$ sebagai berikut:

$$\begin{aligned}
 \mu_{I1}(100) &= \frac{\mu_{100}(I1)}{\mu_{100}(I1) + \mu_{100}(I2) + \mu_{100}(I3)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{3} + \frac{1}{3}} = \frac{\frac{1}{3}}{1} = \frac{1}{3} \\
 \mu_{I1}(400) &= \frac{\mu_{400}(I1)}{\mu_{400}(I1) + \mu_{400}(I2) + \mu_{400}(I3)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{3} + \frac{1}{3}} = \frac{\frac{1}{3}}{1} = \frac{1}{3}
 \end{aligned}$$

Berdasarkan perhitungan tersebut, *item* I1, I2, dan I3 dan dapat didefinisikan sebagai berikut:

$$\begin{aligned}\mu_{11} &= \{(1/3)/100, (1/3)/400, (1/2)/500, (1/2)/700, (1/4)/800, (1/3)/900\} \\ \mu_{12} &= \{(1/3)/100, (1/2)/200, (1/2)/300, (1/3)/400, (1/2)/600, (1/4)/800, (1/3)/900\} \\ \mu_{13} &= \{(1/2)/300, (1/2)/500, (1/2)/600, (1/2)/700, (1/4)/800, (1/3)/900\}\end{aligned}$$

Dari *item-item* tersebut kemudian dicari *confidence* dari setiap kombinasi *k-item* yang memungkinkan, dimulai dari $k = 2$.

Kombinasi 2-item:

$$\begin{aligned}\text{confidence } (I1 \rightarrow I2) &= \frac{\frac{1}{3} + \frac{1}{3} + \frac{1}{4} + \frac{1}{3}}{\frac{9}{4}} = \frac{\frac{5}{4}}{\frac{9}{4}} = \frac{5}{9} = 0,56 = 56\% \\ \text{confidence } (I2 \rightarrow I1) &= \frac{\frac{1}{3} + \frac{1}{3} + \frac{1}{4} + \frac{1}{3}}{\frac{11}{4}} = \frac{\frac{5}{4}}{\frac{11}{4}} = \frac{5}{11} = 0,455 = 45,5\% \\ \text{confidence } (I1 \rightarrow I3) &= \frac{\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}}{\frac{9}{4}} = \frac{\frac{19}{12}}{\frac{9}{4}} = \frac{19}{27} = 0,704 = 70,4\% \\ \text{confidence } (I3 \rightarrow I1) &= \frac{\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}}{\frac{31}{12}} = \frac{\frac{19}{12}}{\frac{31}{12}} = \frac{19}{31} = 0,613 = 61,3\% \\ \text{confidence } (I2 \rightarrow I3) &= \frac{\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}}{\frac{11}{4}} = \frac{\frac{19}{12}}{\frac{11}{4}} = \frac{19}{33} = 0,576 = 57,6\% \\ \text{confidence } (I3 \rightarrow I2) &= \frac{\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}}{\frac{31}{12}} = \frac{\frac{19}{12}}{\frac{31}{12}} = \frac{19}{31} = 0,613 = 61,3\%\end{aligned}$$

Kombinasi 3-item:

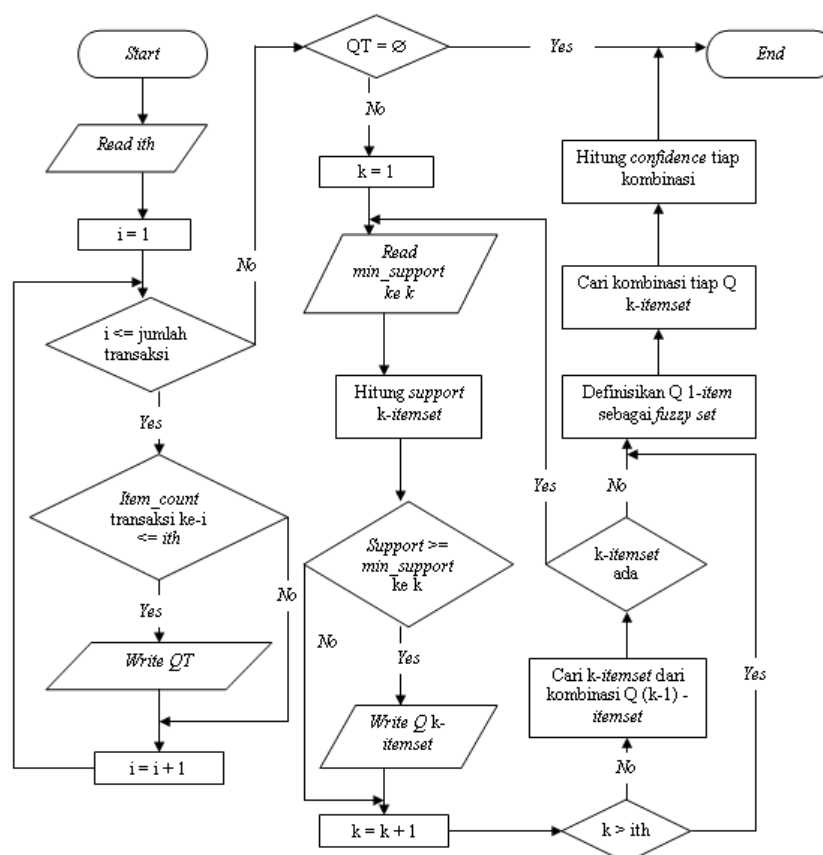
$$\begin{aligned}\text{confidence } (I1 \wedge I2 \rightarrow I3) &= \frac{\frac{1}{4} + \frac{1}{3}}{\frac{1}{3} + \frac{1}{3} + \frac{1}{4} + \frac{1}{3}} = \frac{\frac{7}{12}}{\frac{5}{4}} = \frac{7}{15} = 0,47 = 47\% \\ \text{confidence } (I1 \wedge I3 \rightarrow I2) &= \frac{\frac{1}{4} + \frac{1}{3}}{\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}} = \frac{\frac{7}{12}}{\frac{19}{12}} = \frac{7}{19} = 0,368 = 36,8\% \\ \text{confidence } (I2 \wedge I3 \rightarrow I1) &= \frac{\frac{1}{4} + \frac{1}{3}}{\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}} = \frac{\frac{7}{12}}{\frac{19}{12}} = \frac{7}{19} = 0,368 = 36,8\% \\ \text{confidence } (I1 \rightarrow I2 \wedge I3) &= \frac{\frac{1}{4} + \frac{1}{3}}{\frac{1}{3} + \frac{1}{3} + \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}} = \frac{\frac{7}{12}}{\frac{9}{4}} = \frac{7}{27} = 0,259 = 25,9\% \\ \text{confidence } (I2 \rightarrow I1 \wedge I3) &= \frac{\frac{1}{4} + \frac{1}{3}}{\frac{1}{3} + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}} = \frac{\frac{7}{12}}{\frac{11}{4}} = \frac{7}{33} = 0,21 = 21\% \\ \text{confidence } (I3 \rightarrow I1 \wedge I2) &= \frac{\frac{1}{4} + \frac{1}{3}}{\frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3}} = \frac{\frac{7}{12}}{\frac{31}{12}} = \frac{7}{31} = 0,226 = 22,6\%\end{aligned}$$

Nilai *confidence* dipakai untuk menentukan *rule* mana saja yang merupakan *interesting rule*. Misalkan *user* menentukan *min_confidence* = 60%, maka yang disebut sebagai *interesting rule* hanya *rule* yang nilai *confidence*-nya $\geq 60\%$, yaitu:

if I1 *then* I3 [*support* = 28%, *confidence* = 70,4%]
if I3 *then* I1 [*support* = 28%, *confidence* = 61,3%]
if I3 *then* I2 [*support* = 28%, *confidence* = 61,3%]

Analisis Masalah dan Solusinya

Masalah yang sering dihadapi oleh suatu perusahaan retail yaitu jumlah persediaan barang yang tidak sesuai dengan jumlah permintaan konsumen. Hal tersebut akhirnya menimbulkan permasalahan lainnya, yaitu kapasitas gudang akan terus berkurang karena barang-barang yang tidak laku terjual dan akhirnya harus mengeluarkan biaya tambahan untuk peminjaman gudang, konsumen akan pergi mencari perusahaan lain jika persediaan barang di perusahaan tersebut tidak ada atau tidak mencukupi. Salah satu cara untuk meminimalisasi permasalahan tersebut adalah dengan *Market Basket Analysis*, yang berguna untuk menganalisis data penjualan dari suatu perusahaan, dan hasil yang didapatkan ini nantinya dapat digunakan untuk mengembangkan strategi pemasaran dengan melihat barang-barang mana saja yang sering dibeli oleh konsumen. Setelah melihat permasalahan di atas dan mempelajarinya, penulis mencoba untuk merancang suatu program aplikasi yang menggunakan *Market Basket Analysis* dan logika fuzzy untuk menyelesaikan permasalahan tersebut. Logika fuzzy yang akan digunakan adalah *Fuzzy c-Covering*. Berikut ini adalah *flowchart* algoritma untuk aplikasi *Market Basket Analysis* berdasarkan metode *Fuzzy c-Covering* (Gambar 1).



Gambar 1. Flowchart algoritma market basket analysis.

PENUTUP

Setelah melakukan analisis dan perancangan program aplikasi *Market Basket Analysis*, dapat diperoleh kesimpulan sebagai berikut; (1) memberikan hasil analisis kepada perusahaan untuk memperkirakan persediaan barang pada masa yang akan datang; (2) nilai *support* dan *confidence* merupakan bagian dari *Market Basket Analysis* yang dihitung dengan menggunakan metode *Fuzzy c-Covering*. *User* dapat memberikan nilai batasan *minimum support* dan *minimum confidence* sesuai dengan keinginannya sendiri. Semakin tinggi nilai batasan, hasil analisis yang didapat akan lebih terseleksi; (3) data yang digunakan adalah data transaksi penjualan satu bulan terakhir. Jika data yang digunakan lebih banyak, hasil analisis yang didapat akan lebih akurat. Oleh karena itu, tingkat keragaman data sangat mempengaruhi *Market Basket Analysis*.

Beberapa saran yang diajukan dengan kemungkinan dilakukan pengembangan lebih lanjut adalah sebagai berikut; (1) pengembangan lebih lanjut disarankan untuk dapat menggunakan metode *Fuzzy c-Covering* pada penerapan aplikasi lainnya; (2) pengembangan lebih lanjut disarankan untuk dapat menggunakan *Market Basket Analysis* pada penerapan aplikasi lainnya; (3) pengembangan lebih lanjut dapat diterapkan pada sistem *database* yang menggunakan *server*, sehingga data yang digunakan merupakan data yang terbaru, jadi hasil analisis yang didapat akan lebih akurat.

DAFTAR PUSTAKA

- Albion Research Ltd. (2007). *Market Basket Analysis*. Diakses November 2009 dari http://www.albionresearch.com/data_mining/market_basket.php.
- Han, Jiawei & Kamber, Micheline. (2001). *Data Mining: Concepts and Techniques*. Massachusetts: Morgan Kaufmann.
- Hoobs, L., Hillson, S. & Lawande, S. (2003). *Oracle 9iR2 Data Warehousing*. Burlington: Digital Press.
- Klir, George J. & Yuan, Bo. (2001). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. New Delhi: Prentice Hall.
- Megaputer. (2007). *Market Basket Analysis*. Diakses Desember 2009 dari http://www.megaputer.com/company/cases/cambridge_mba.php3
- Ter Meulen, Alice. (2006) *Logic and Natural Language*. Diakses Desember 2009 dari <http://en.wikipedia.org/wiki/Naturallanguage>
- Ulmer, David. (2002). Mining an online auctions data warehouse. *The Mid-Atlantic Student Workshop on Programming Languages and Systems*. Pace University, New York.